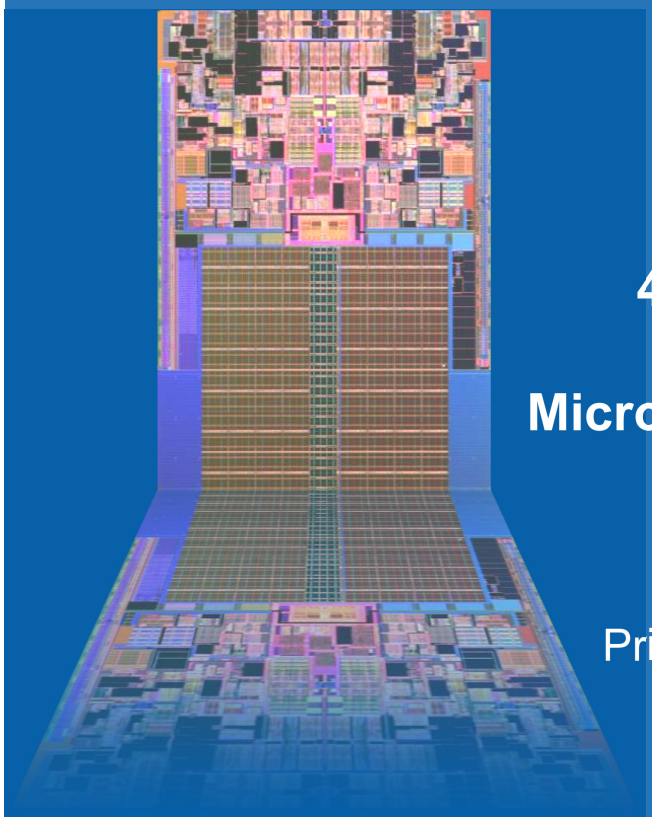


45nm Next Generation Intel® Core™ Microarchitecture (Penryn)

HOT CHIPS 2007

Varghese George
Principal Engineer, Intel Corp

Legal Disclaimer

Outline

Next Generation 45nm Process Technology benefits

Overview of Penryn Architecture and New Features

Highlights of New Microarchitecture changes

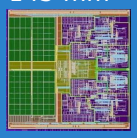
New Power Management Features in Penryn

- Deep Power Down Technology
- Enhanced Dynamic Acceleration Technology
- CC3 in Servers

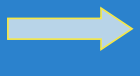
Summary

Next generation 45 nm Hi-K metal gate process technology: Continuing Moore's Law

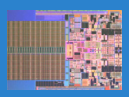
65 nm
Intel® Core™ 2 Duo
143 mm²



4-wide Dynamic Execution
Advanced smart cache: 4M L2
Smart Memory access
Advanced digital media boost
Intelligent power capability



45 nm
Penryn family die (dual core)
107 mm²



Same good stuff plus:
- 6M L2 cache
- 47 new SSE4 instructions
- Micro-arch enhancements for further IPC perf
- Deep Power Down Technology*
- Enhanced Dynamic Acceleration Technology*

Penryn advantages with this next generation process:

- Lower transistor switching power and leakage current => reduced idle power => longer battery life, quieter systems
- Up to 2x improvement in transistor density: Significant area scalability to support additional capabilities within the existing power envelope
- Higher core and bus frequency for more performance in the same power envelope

** Mobile only features*

Enhanced Intel® Core™ Microarchitecture

Intel Core Microarchitecture	*New with the Penryn Family*
Intel® Wide Dynamic Execution	Fast Radix-16 Divider Faster OS Primitive Support Enhanced Intel Virtualization Technology
Intel® Advanced Smart Cache	Larger Caches: up to 12MB 24 Way Set Associativity
Intel® Smart Memory Access	Improved Store Forwarding Higher bus speeds
Intel® Advanced Digital Media Boost	Intel SSE4 instructions Super Shuffle Engine
Intel® Intelligent Power Capability	Deep Power Down Technology Enhanced Intel Dynamic Acceleration Tech

Increased Performance and Energy Efficiency Across Applications

Covered briefly in this presentation

Focus of this presentation

New Intel® SSE4 Instructions

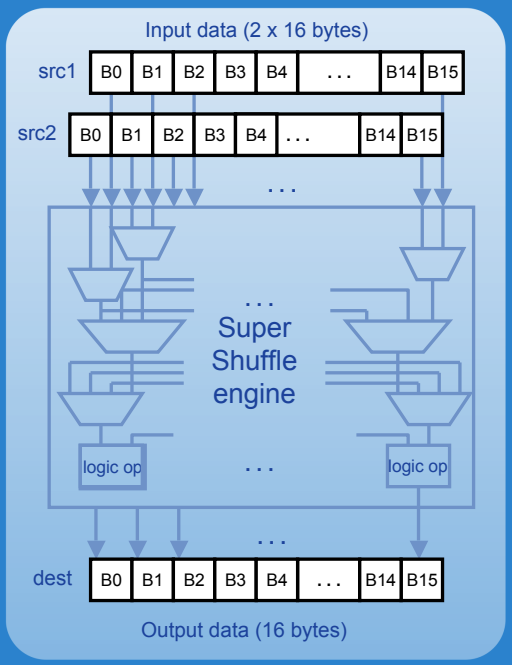
Intel® SSE 1999	Intel® SSE2 2000	Intel® SSE3 2004	Intel® SSE3 2006	Intel® SSE4 2007
70 instructions •Single-Precision Vectors, •streaming operations	144 instructions •Double-precision Vectors •128-bit vector integer	13 instructions •Complex Arithmetic	32 instructions • Decode	**47 instructions** **• Video Accelerators** **• Graphics building blocks** **• Coprocessor Accelerators**

- Enhanced Intel® Core™ Microarchitecture (Penryn) supports 47 new instructions
- This continues a trend set by SSE, SSE2, SSE3, and SSSE3

SSE4

- Largest set of new instructions introduced since 2000
  - Focused on further performance gains for SIMD software via SSE
- Addresses some popular requests for key functionality
  - Enhances SSE/2 orthogonality, support for vectored 32-bit operations
- Features to improve the compiler's ability to efficiently vectorize code
  - Blends, tests, rounds, and sign/zero extensions are straightforward replacements for existing lengthy operations
  - Inserts, extracts are building blocks to gathers (lookups), scatters, strided loads, and strided stores
- Also adds highly specialized operations
  - Video Encode acceleration functions
  - Floating-point dot product operation (i.e. 3D content)
  - Streaming load for high b/w to WC memory (imaging, GPU-CPU sharing)

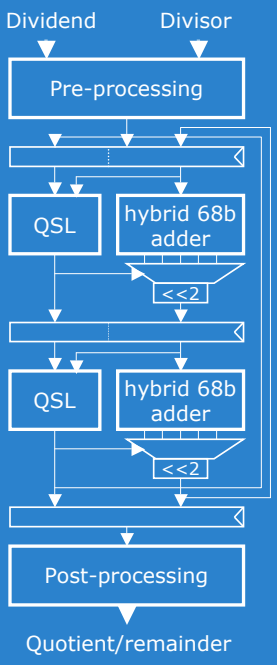
Super Shuffle Engine



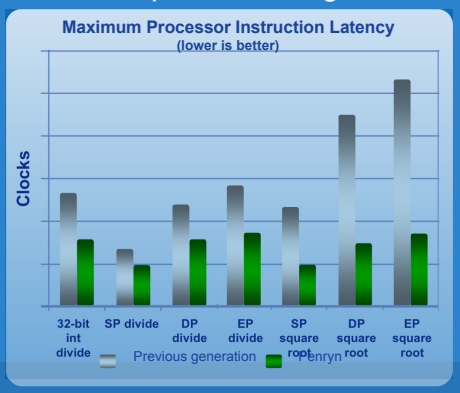
- Shuffle operations required for SSE data formatting operations
  - Unpacking
  - Packing
  - Align concatenated sources
  - Wide shifts
  - Insertion and extraction
  - Setup for horizontal arithmetic functions
- Penryn super shuffle engine performs 128 bit operation in a single cycle
- No software changes required

Doubles shuffle throughput

New Radix-16 divider architecture



- Leading edge divider performance
  - >50% speed-up over previous generations on average
- Innovative radix-16 based architecture
  - Double pumped: processes 2x2 bits per cycle
  - Hybrid 68-bit CSA/CPA based for faster pipelined quotient select
  - Optimized square root
- Minimized data dependent integer & FP divide latency



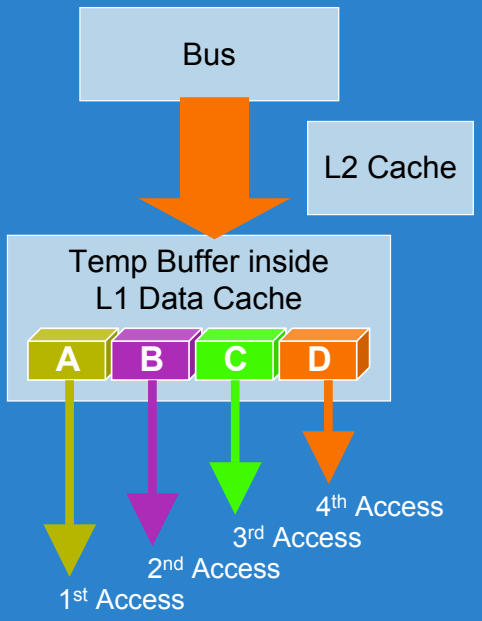
Streaming Load Instruction

16 byte Aligned load instruction on WC (write combining) memory with non-temporal hint (compare to streaming store)

- Improves read bandwidth from WC mem by reading cache-line size quantities
- Results held in a temporary stream buffer
- Cache is not polluted
- Buffered data is intended for single use
- Example usage:

```
MOVNTDQA xmm0, [eax]
MOVNTDQA xmm1, [eax+16]
MOVNTDQA xmm2, [eax+32]
MOVNTDQA xmm3, [eax+48]
PAVGB xmm0, xmm1
PAVGB xmm2, xmm3
PAVGB xmm0, xmm2
```

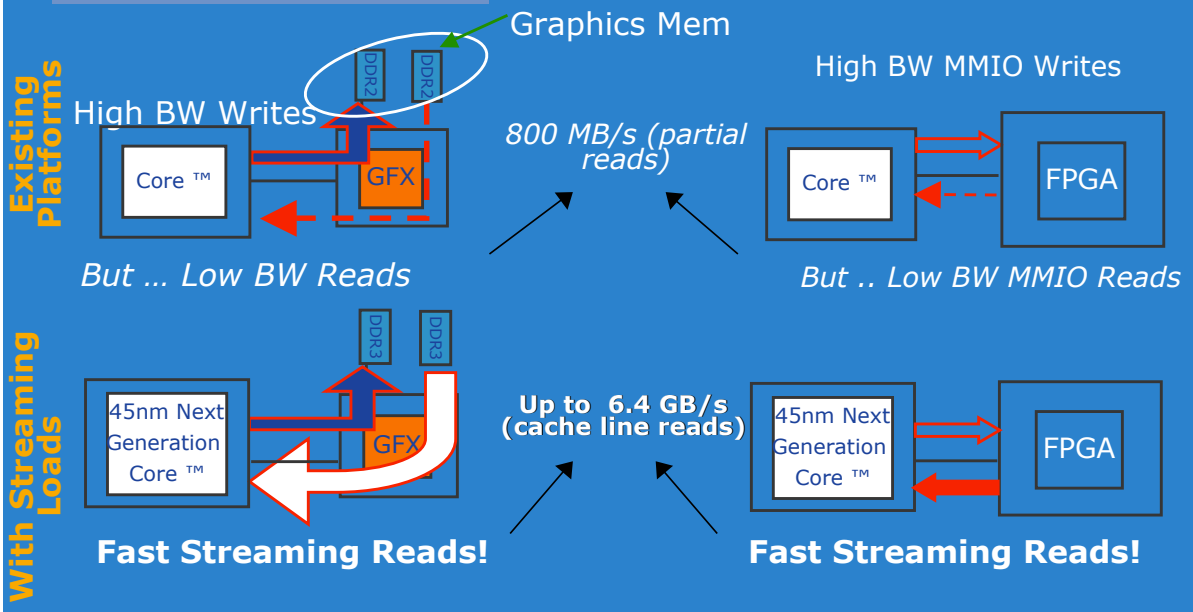
For maximum perf, the s/w should issue 4 loads closely together to read the full buffer



(Data usage in processor core)

Streaming Load Bandwidth improvement

MOVNTDQA xmm, m128



Streaming Load is ~8X faster reading from WC Memory

Power Management Overview

Penryn builds on the capabilities in the Core™ microarchitecture

- Fine grained clock gating during runtime: Applied all across the chip
- Sleep transistors: In large arrays such as L2 cache to reduce leakage power
- C-states: Various low power states during CPU idle
- P-states: Multiple performance states for run-time energy efficiency
- Thermal Monitor: Self-managed thermal control based on multiple thermal sensors
- Digital Thermal Sensor: For software access to CPU temperature
- Intelligent Voltage Regulator control: To optimize VR efficiency during low loads
- PECI interface: For side-band access to temperature for fan control

Key New Mobile Power/Thermal Features in Penryn

- **Deep Power Down Technology (DPD):** *A radically new and innovative idle power management state*
- **Enhanced Dynamic Acceleration Technology (EDAT):** *A deterministic and (TDP) power-neutral performance boost for Single Threaded apps*

Deep Power Down Technology Overview

Concept: When idle, required CPU state is saved in a power-preserved domain on chip and then power is shut off

Exit: Upon a break event (such as interrupt), Power and Clocks are turned on, CPU is Reset, State is restored and execution resumes seamlessly from where it left off transparent to OS / applications

- Can occur few 100s of times per second (between keystrokes)

Salient features:

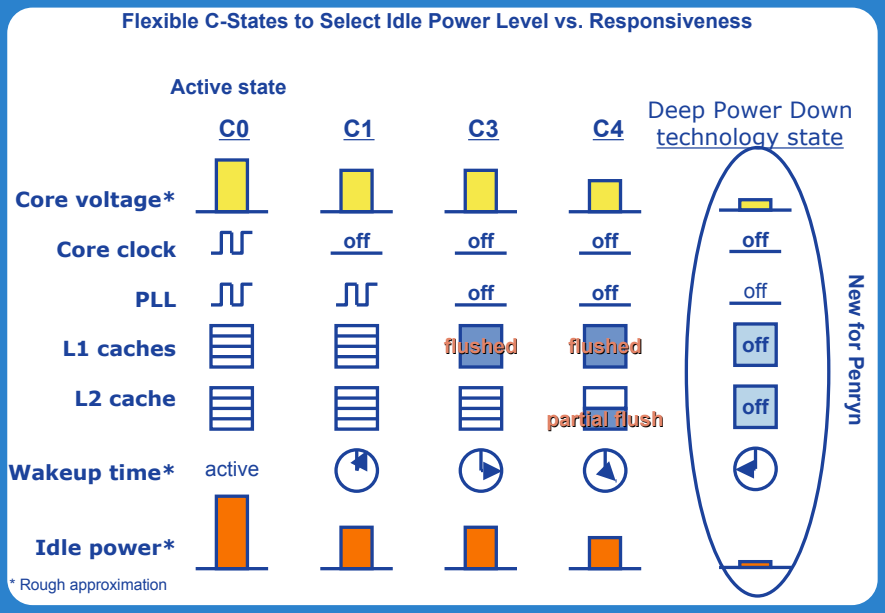
- Caches sharing power plane with core need to be flushed
- No Software intervention necessary for context save/restore
- Exit latency well within today's C-state exit latency bounds (few hundred microseconds)
- No CPU wakeup necessary for snoop traffic

Key Power benefits of DPD:

- Lowest Leakage CPU state you can get to!
- Agnostic to min Vcc state retention issues

DPD enables reaching lower limit of CPU idle power of 0 W

DPD vs. Other C-states



State in DPD

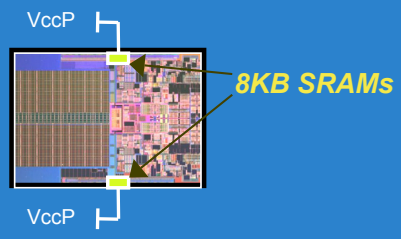
Vcc: Well below retention

Core clk: OFF

PLL: OFF

Caches: OFF

Penryn DPD Implementation Overview



4 Major parts:

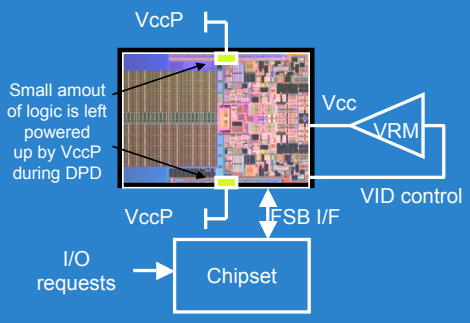
STATE STORAGE:
- 8KB (per core)
- Powered from I/O Vcc (VccP)
- ECC protected

MICROCODE:
- Does state save and restore
- Manages core synchronization
- Supports Auto-demote policy

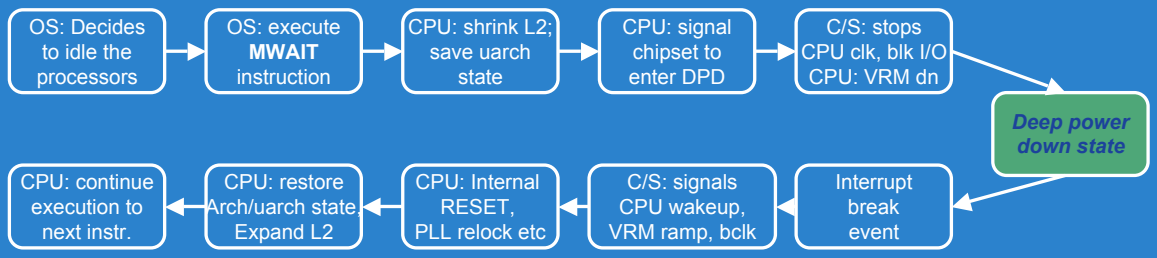
PM UNIT ("hardware"):
- Coordinates cores at hardware level
- Manages the DPD Power up sequence
- Manages entry/exit protocol with platform

STATE DEFINITION:
- What to include?
- Criteria: *"Software seamless"*
- Inclusions:
  - All Architectural state
  - Most micro-architectural state
- Exclusions:
  - Temp registers used by ucode
  - Some others on a case by case basis

Deep Power Down Technology Entry/Exit



- S/W instruction initiates processor DPD entry
- CPU does rest of sequencing with chipset/platform
- Protocol with chipset to block snoops due to DMA traffic (no CPU wakeup required) while in DPD state
- Exit initiated by an break event in platform (interrupt) through the chipset
- CPU sequences through rest of exit sequence – including driving VID to VRM, internal hardware reset, state restore and execution resumption



Interrupt Rate Sensitivity of Average Power

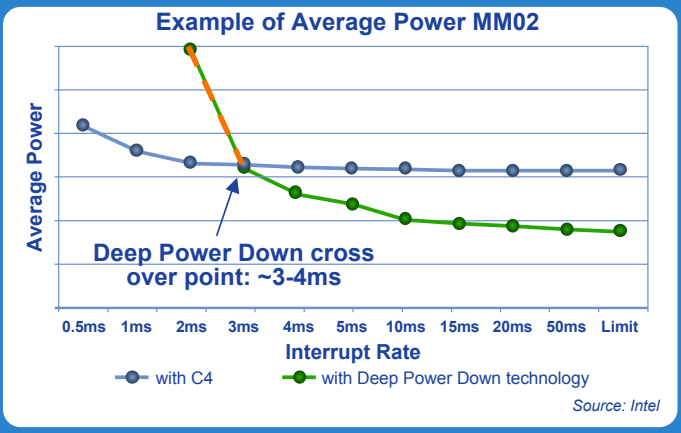
Deeper C-states have non-trivial energy cost for transitions

Today, OS makes C-state decisions based on latency, not energy cost

- Latency could be ok, but energy cost could be significantly different

Too frequent transitions into deep C-states could result in net energy loss

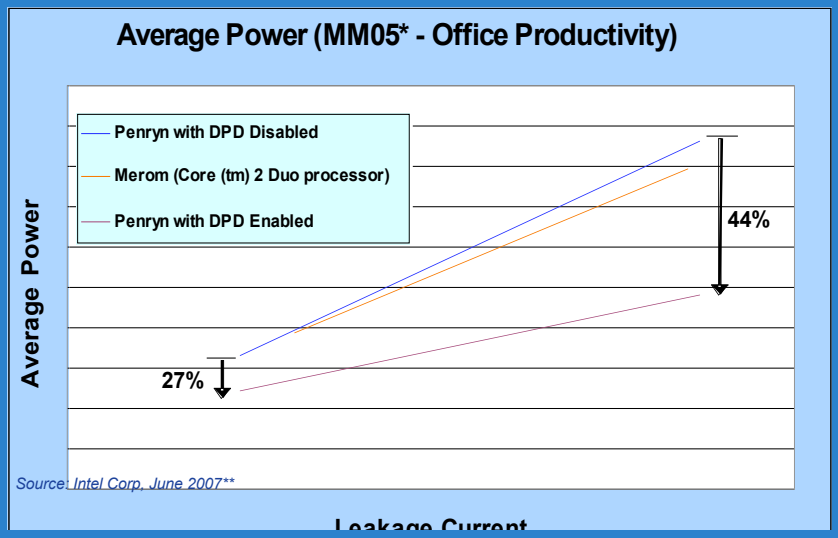
- Worse battery life



Penryn implements a *"Auto-demote"* policy in the CPU to address this issue

- Heuristics to detects when energy cost of DPD state entry and exit is potentially > savings in DPD state
- Demotes such cases to shallower C-state (C4), although the OS requested DPD
- Silicon results prove that this was a good decision!

DPD Results (Average Power)



Actual Results measured in Penryn Silicon – Up to 44% reduction in Average Power

- 27% to 44% (based on the leakage of the part) reduction on Average Power as measured by Mobile Mark – Office Productivity benchmark due to DPD feature
- Significant improvement compared to previous generation (Merom)
- Measured Exit latency for DPD state: ~ 150 - 200 us => In expected range

***Config: Data from a sample distribution of parts with various leakage measured on tester, under respective product POR Vcc, temperature etc*

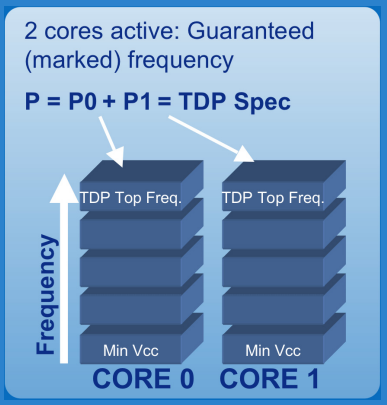
MM05 – BAPCO Mobile Mark* 2005 benchmark*

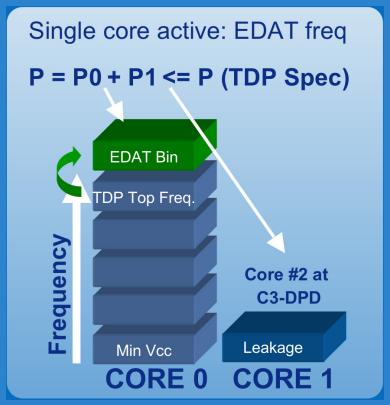
Enhanced Dynamic Acceleration Technology (EDAT)

Concept: In multi-core CPUs, use the power headroom of idle core to boost performance of the non-idle core

How it works:

- When one core enters an idle power C-state (CC3 or deeper) AND
- OS requests a higher performance state on the running core,
- The non-idle core is boosted up to a higher voltage, higher frequency (EDAT freq)
- Overall chip Power still remains within the specified Thermal Design Power (TDP)

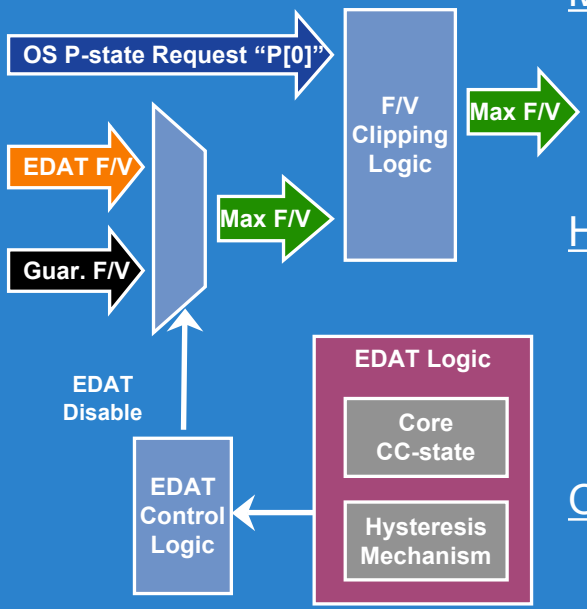




Applicability: EDAT is applicable mostly to thermally constrained platforms such as mobile where the guaranteed frequency is not max Vcc limited

EDAT provides a significant and predictable single-threaded performance boost

EDAT Implementation Overview



Microarchitecture

- Entry based on OS request AND other core idle
- Idle core defined as "CC3" or deeper C-state
- EDAT Freq pre-programmed in chip based on power, reliability and other constraints
- Exit EDAT mode when Idle core wakes up

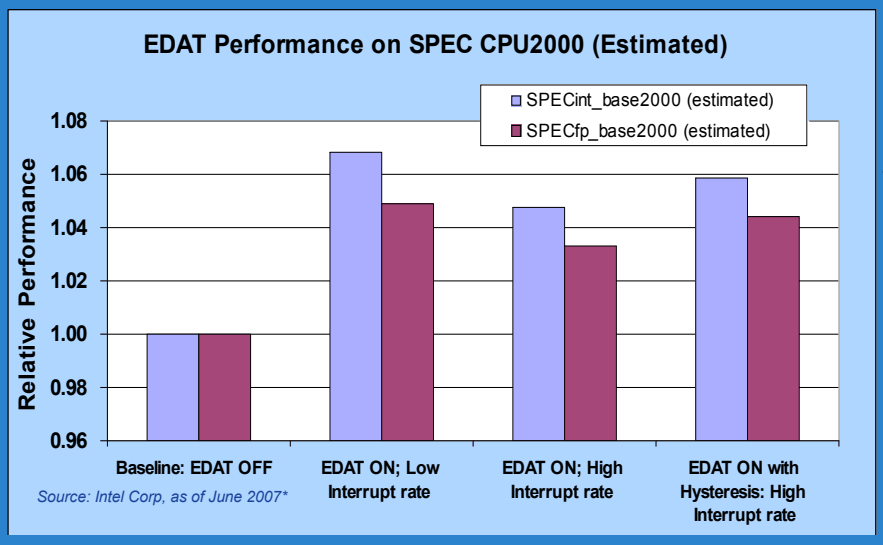
Hysteresis mechanism

- Allows short durations where 2 cores active
- Reduces perf loss for low activity wakeups
- Implemented using a few counters
- Voltage Regulator needs to provide for this
- Benefits most at high timer tick rates

OS interface

- EDAT Freq exposed as {Guaranteed + epsilon} P-state in ACPI table: P[0]
- OS requests P[0] state if perf demand exists
- EDAT logic grants it if power headroom exists

EDAT Results (Performance)



Actual Results measured in Penryn Silicon – Up to 7% boost in ST perf due to EDAT

- 5% (on SPECfp_base2000 (est*)) and 7% (on SPECint_base2000 (est*)) performance gains due to EDAT within the same TDP power envelope!**
- High timer tick rates would have reduced the gains by a few % due to frequent wake up of idle core and resulting exits from EDAT mode
- Intelligent hysteresis mechanism in Penryn recovers most of that loss

**Configuration used for measurement: Matanzas CRB board, Crestline chipset DDR2-667 memory, Penryn A1 silicon @ 2.4 GHz, FSB 800, OS: Windows XP SP2

*Estimated SPEC metrics based on measurements on preproduction platforms. SPEC, SPECint and SPECfp are trademarks of SPEC. For more information on these benchmarks, see www.spec.org

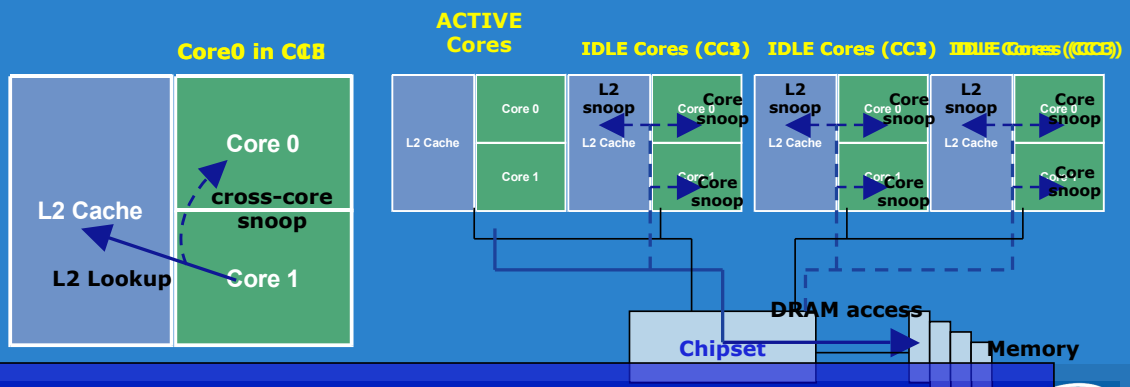
New PM feature for Servers: CC3 State

Concept:

- Snoops burn ~30% of active core power
- By avoiding snoops into idle cores, we save power

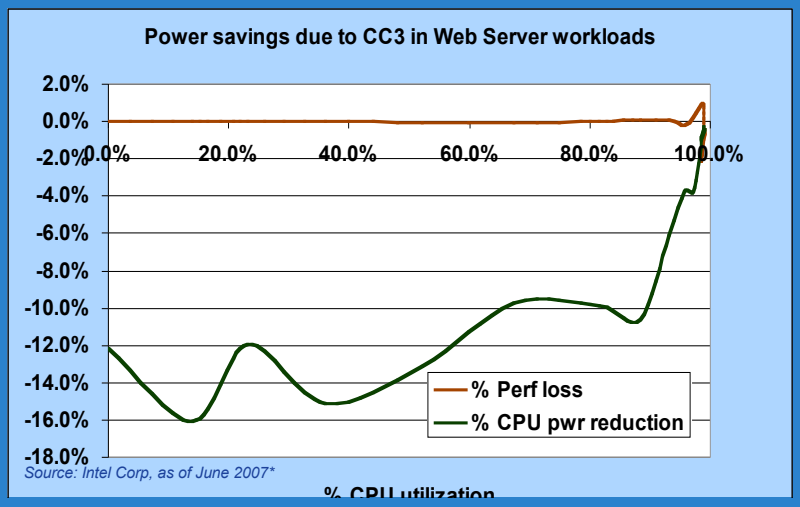
How it works:

- Today, idle cores are put into Core C1 (CC1), which is a snoopable state
- In Penryn, idle cores can be put into Core C3 (CC3), which is a non-snoopable state.
- First level caches are flushed into the L2 cache before putting cores into CC3
- This prevents cross core snoops and therefore the additional power burnt for snoops



CC3 is a simple but effective optimization for power savings

CC3 Results (Power savings)



Source: Intel Corp, as of June 2007*

Up to 16% CPU power reduction measured on web server workloads with CC3 feature

- Up to ~16%** CPU power reduction measured on Web Server workload with negligible performance loss with CC3 feature enabled
- Similar power reductions seen on SPEC CPU2000 (estimated by measurement on preproduction silicon) (single user mode), other idle power benchmarks for Server

*Configuration used for measurement: Bridgeport2, DP Clovertown 2.33 GHz, 1333 MHz Bus, 4x2GB FB DIMM 667 MHz, OS: Linux 2.6.18.1; Workload: Specweb99; DBS OFF; Measured by mapping CC3 as CC1 in BIOS

** Note that the power savings due to this feature will be system configuration, OS and workload dependent.

Summary

New generation Intel® Core™ microarchitecture is leading the way with 45 nm process technology

New microarchitecture enhancements for higher performance and improved energy efficiency

New SSE4 instructions for improved video, imaging, and 3D content performance

New Power management features for improved battery life, performance and energy efficiency

Moore's Law is alive and well with Penryn; Innovative features with tangible benefits being introduced for Performance and Power

Acknowledgements

To the entire Penryn team across Folsom, Santa Clara, Costa Rica and Israel and the Technology Development Group for developing and delivering this product to market

Legal Notices and Important Information

Regarding the performance and power measurements in this presentation